

국외 출장 결과보고서

결	담당자	팀장	실장	사업단장
	12월 9일	12월 9일	12월 9일	12월 10일
재	노일석	조희제	김정한	박훈

1. 출장개요

출장목적	SC 2025 (The International Conference for High Performance Computing, Networking, Storage, and Analysis 2025) 참석		
출장일정	2025. 11. 15. ~ 2025. 11. 23. (7박 9일)		
성명		소속	직급(직위)
노일석		모델체계팀	선임급연구원
방문기관 (또는 회의장소)	아메리카스 센터 (America's Center, St. Louis, Missouri, United States of America)		
출발지	인천(Incheon)	도착지	세인트루이스(St. Louis)

2. 업무수행내용

가. 출장 배경 및 목적

○ 개요

2025년 11월 16일부터 21일까지 미국 미주리주 세인트루이스에서 열리는 The International Conference for High Performance Computing, Networking, Storage, and Analysis 2025 (이하, SC25)은 사업단에서 연구/개발/운영하는 수치예보모델의 구동에 필수적인 HPC 시스템을 주제로 전 세계 각 나라의 시스템과 엔지니어가 모여 의견을 나누며 보고 배울 수 있는 만남의 장소이다. 사업단의 KIM 모델은 거대한 자료를 기반으로 방대한 계산자원을 요구하기 때문에 HPC 및 Cluster 시스템은 KIM 모델 개발/운영에 필수적인 요소이며 이를 최적의 환경으로 구성하는 것은 KIM 모델 개발/운영에 중요한 축이다.

특히, 이번 SC25는 지난해에 이어 대규모 GPU 시스템을 활용한 AI 활용 기술, 새로운 양자컴퓨팅 기술 등이 중점적으로 소개될 예정이다.

○ 목적

현업 KIM 수행시간에서 상당한 부분을 KIM 계산 결과를 파일로 저장하는 과정이 차지하고 있다. 예측 해상도가 상세해질수록 입출력에 필요한 시간은 증가함으로 반드시 최적화가 필요한 부분이다. SC25에서 거대 배열의 병렬통신과 대용량 입출력 파일을 다루는 기술을 위주로 해외 전문가들의

연구 경험과 기술을 파악하고 KIM 최적화에 관련 기술들을 도입하고자 한다. 그리고, 최근 일기 예보 분야에서 AI 기술을 활용한 날씨예측모델이 빠르게 개발되고 있다. SC25에서는 대규모 GPU 시스템을 사용한 거대 AI 모델 관련 연구가 소개될 예정이며 관련 동향을 파악하고자 한다.

나. 주요 업무 수행 사항

○ 수치예보모델 및 프레임워크 최적화 관련 연구

- 고해상도 수치 모델링 및 거대 기상·기후 AI 학습이 주도하는 현재 수치 모델링 연구 환경에서는 수십 PB 크기의 데이터를 하루에 처리하게 된다. 데이터를 압축해 파일 크기를 줄이는 접근은 효율적으로 입출력 시스템과 데이터센터의 저장공간을 관리하는 방안이며 데이터의 크기가 계속해서 커지고 있는 현재 환경에서는 필수적이다. 스위스의 Swiss National Supercomputing Centre(CSCS)에서는 그래픽 압축 기술 JPEG을 수치모델 데이터 처리에 활용한 기법을 소개했다. 이 기법의 핵심은 데이터가 압축될 때, 일부 픽셀(격자점)의 정보는 허용할 수 없는 오차 경계 (error bound)를 넘어서게 되는데 그 지점들에 대한 압축을 반복하며 허용할 수 있는 수준으로 오차로 만드는 것이다. 여기서 허용할 수 있는 오차 경계는 현업 앙상블 예측에서 앙상블 멤버 간 RMSE로 결정한다.
- 독일의 Max Planck Institute for Meteorology에서는 Icosahedral Nonhydrostatic(ICON) 모델을 대기, 지면, 해양, 식생, 해양생화학 모델을 결합한 1km 해상도 지구 시뮬레이션 결과를 발표하였다. 특히, 식생 모델과 해양생화학 모델을 결합해 지구 에너지와 탄소 순환 연구에 활용할 수 있음을 보였다. 독일의 슈퍼컴퓨터 JUPITER 시스템의 85% 수준인 GH200 20,480개를 사용해 24시간 계산에 147.7일을 계산하는 속도를 기록했다. 이 연구에서 대기, 지면, 식생 모델은 GB200 칩의 GPU 영역에서, 해양, 해양생화학 모델은 GB200 칩의 CPU 영역에서 계산했으며, 포트란 코드를 C++ 기반 코드로 변환하는 Data-Centric Parallel Programming(DaCe) 프레임워크를 적용하였다.
- 일본의 이화학연구소(RIKEN)은 오사카 엑스포 일기예보를 위한 30초 간격으로 갱신되는 500m 수평해상도의 초고해상도 예보를 한 달간 수행한 경험을 발표하였다. 이 예보모델은 Local Ensemble Transform Kalman Filter(LETKF) 1000개 멤버를 사용한 레이더 자료동화를 바탕으로 11개 멤버의 앙상블 지역 모델로 일본의 슈퍼컴퓨터 Fugaku 시스템의 16%를 사용하였다.

3. 출장성과 · 시사점 및 향후 업무 활용 계획

가. 출장성과 · 시사점

- 사업단의 중기, 단기, 초단기 예측체계 산출물의 크기가 증가하며 KIM에서는 입출력의 부하 문제, 기상청 슈퍼컴퓨터는 저장공간 관리의 문제를 경험하고 있다. 이러한 문제는 국내 계산 환경보다 더 큰 시스템을 갖춘 해외 기관에서도 중요한 문제로 인식해 대응 중인 것은 데이터 압축 기술에 대한 중요도를 시사한다.
- 최근 몇 년간 포트란 기반 수치해석 프로그램들의 GPU 이식 과정에서 지시어 기반 라이브러리보다 프로그래밍 언어를 C++로 변경해 재작성하는 방식을 채택한 연구들이 좋은 성과를 보여주었다. 특히, 컴퓨터 과학 분야에서 C++ 프로그래머로 인력이 집중되며 포트란의 GPU 연산 환경이 C++ 환경만큼 GPU 칩셋 발전 속도에 대응하지 못하고 있다.

나. 향후 업무 활용 계획

- KIM 프레임워크에서 현재 사용하고 있는 NetCDF 압축 외에 학회에서 소개된 다양한 압축기법을 적용할 수 있는지 검토할 계획이며, ICON에서 사용한 DaCe 프레임워크를 현재 KIM 코드에서 활용할 수 있는지 분석하고자 한다.